

**PATENT**  
Attorney Docket 2676-4555.1US

CERTIFICATE OF MAILING

Express Mail Label Number: EV325783810US

Date of Deposit: January 15, 2004

Person making Deposit: Christopher Haughton

PATENT APPLICATION

for

**CD40-INTERACTING AND TRAF-INTERACTING PROTEINS**

Inventors:

Stefan M. C. Pype  
Jacques E. F. Remacle  
Danny F. E. Huylebroeck

Attorney:  
Allen C. Turner  
Registration No. 33,041  
TRASK BRITT, PC  
PO Box 2550  
Salt Lake City, Utah 84110  
(801) 532-1922

## CD40-INTERACTING AND TRAF-INTERACTING PROTEINS

### CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application is a divisional application of co-pending application U.S.S.N. 09/687,863, filed October 27, 2000, now U.S. Patent \_\_\_\_\_, which is a continuation of PCT International Application No. PCT/EP99/03025, filed on April 28, 1999, designating the United States of America, and published in English as WO 99/55859 on November 4, 1999, the contents of both of which are incorporated by this reference.

### TECHNICAL FIELD

[0002] The invention relates to CD40 binding proteins, which can be used as modulators of the CD40 signaling pathway and/or the CD40-induced nuclear factor kappa B (NF- $\kappa$ B) activating pathway and thus useful in the treatment of CD40 related diseases (*e.g.*, inflammatory diseases) and/or NF- $\kappa$ B related diseases and/or in the improvement of anti-tumor treatments. The current invention also relates to nucleic acid sequences coding for the CD40 interacting proteins (also called "TTRAP" ("TRAF and TNF receptor associated protein") for CD40 receptor associated protein). The invention further relates to the use of the polypeptides derived from these CD40 interacting proteins in the treatment of CD40 and/or NF- $\kappa$ B related diseases and/or cancer. Furthermore, the invention concerns pharmaceutical preparations comprising the CD40 interacting proteins or polypeptides derived from these proteins.

### BACKGROUND

[0003] CD40 is a receptor of the tumor necrosis factor ("TNF") - receptor superfamily (Banchereau *et al.*, 1994), which is expressed at the surface of B-cells, antigen presenting cells (APC), and several non-hematopoietic cells such as endothelial cells (Hollenbaugh *et al.*, 1995), epithelial cells (Galy & Spits, 1992), fibroblasts (Fries *et al.*, 1995) and keratinocytes (Gaspari *et al.*, 1996). The ligand for CD40 (CD40L) occurs mainly on activated T-cells. Up to now the role of CD40 was mainly studied in the context of the T-cell APC/B-cell interaction (for a review, see Noelle, 1996). Amongst others, the CD40-CD40L interaction seems to be important for the T-cell

mediated immunity and for primary and secondary humoral immune response. These findings were confirmed by experiments in mouse models showing that treatment with anti-CD40L antibodies resulted in blocking of the development of mouse equivalents of human autoimmune diseases such as arthritis (Durie *et al.* 1993), oophoritis (Griggs *et al.*, 1996) and multiple sclerosis (Gerritse *et al.*, 1996).

**[0004]** Activation and transduction through the CD40 pathway within this biological system is mainly responsible for B cell activation and the humoral immune response accordingly.

**[0005]** Apart from NF- $\kappa$ B, factors that can be activated by stimulation of CD40 are NF-AT (Francis *et al.*, 1995) c-Jun, ATF-2 and IRF-1 (Karmann *et al.*, 1996). All these factors play an important role in inflammation.

**[0006]** The CD40L induced signal transduction is, like TNF, mediated by the binding of TNF-Receptor Associated Factors (TRAF's) to the cytoplasmic domain of the receptor. Chaudhuri *et al.* (1997) demonstrated that, at least in human B cell lines, CD40 and TRAF2 are constitutively associated with each other, and that this association is inhibited by CD40 mediated signals. Apart from the binding with TRAF2, the cytoplasmic domain of CD40, which consists of 62 amino acids at positions 196-257 (mature human CD40 - numbering according to Kashiwada *et al.*, 1998), is known to associate with TRAF3, TRAF5, TRAF6 and Janus kinase 3. TRAF6 binds to the amino-terminal cytoplasmic tail of CD40 at positions 210-225, although it can not be excluded that full association of TRAF6 with CD40 may also require the carboxy-terminal part at positions 226-249 (Ishida *et al.*, 1996). TRAF2, TRAF3 and TRAF5 bind to the carboxy-terminal CD40 cytoplasmic domain at positions 226-249 (Ishida *et al.*, 1996).

**[0007]** Stimulation of CD40 results in activation of protein kinases, the mitogen-activated protein kinase and Janus kinase 3 / signal transducer and activator of Transcription 3. Moreover, stimulation of CD40 mediates critical biological effects in B cell growth, survival and differentiation.

**[0008]** It is known that TRAF2 and TRAF5 play a role in NF- $\kappa$ B activation in signaling through CD40, as well as TNF-RI, TNF-RII, CD30 and lymphotoxin b receptor. TRAF6 participates in NF- $\kappa$ B activation signaled by CD40 and IL-1 receptor. In addition to these data, International Patent Applications WO 96/16665 and WO 96/28568 disclose a TRAF like protein that binds to the cytoplasmic domain of CD40.

## DESCRIPTION OF THE INVENTION

**[0009]** We show herein that at least two other proteins exist which unexpectedly interact with the cytoplasmic domain of CD40. Even more surprisingly, neither of these proteins shows significant homology with any known CD40 interacting proteins. Further, no significant homology exists between the two proteins themselves. These proteins should, therefore, be considered as two new classes of CD40 interacting proteins.

**[0010]** The present invention, thus, concerns an isolated functional protein capable of interacting with the cytoplasmic domain of CD40 and/or other receptors of the TNF receptor superfamily such as CD30 or TNF receptor II, wherein the protein has no homology to TRAF-proteins.

**[0011]** The invention also includes an isolated functional protein either comprising an amino acid sequence with 70-100% homology to the amino acid sequence depicted in SEQ ID NO:2 or either comprising an amino acid sequence with 70-100% homology to the amino acid sequence depicted in SEQ ID NO:4 or, in the alternative, comprising an amino acid sequence with 70-100% homology to the amino acid sequence depicted in SEQ ID NO:6.

**[0012]** More specifically, the functional protein comprises an amino acid sequence with 70-100% homology to the amino acids 54-362 of SEQ ID NO:2, even more specifically, the functional protein comprises an amino acid sequence with 70-100% homology to the amino acids 274-362 of SEQ ID NO:2 or in the alternative and/or comprising an amino acid sequence with 70-100% homology to the amino acids 2-245 of SEQ ID NO:6.

**[0013]** Furthermore, the invention also includes those proteins or peptides having 70-100% homology to, at least, any of the three peptides as depicted in SEQ ID NO:2 located between the residues 115-121, 145-153 and 347-352, respectively. The amino acid sequence of residue numbering 115-121 of SEQ ID NO:2 is SLITWNI, the amino acid sequence of residue numbering 145-153 of SEQ ID NO:2 is PDVIFLQEV and the amino acid sequence of residue numbering 347-352 of SEQ ID NO:2 is FPSDHW.

**[0014]** "Homology," in this context, means identical or similar to the referenced sequence while straightforward replacements/modifications of any of the amino acids provided, are included as

well. A homology search in this respect can be performed with the BLAST-P (Basic Local Alignment Search Tool), a program well known to those of skill in the art. For the corresponding nucleic acid sequence, homology is referred to the BLASTX and BLASTN programs known in the art.

**[0015]** The invention thus relates to a DNA sequence encoding a CD40 receptor associated protein or encoding an immunologically active and/or functional fragment of such a protein, selected from the group consisting of:

- [0016]** (a) DNA sequences comprising a nucleotide sequence encoding a protein comprising the amino acid sequence as given in SEQ ID NO:2;
- [0017]** (b) DNA sequences comprising a nucleotide sequence as given in SEQ ID NO:1;
- [0018]** (c) DNA sequences hybridizing with the complementary strand of a DNA sequence as defined in (a) or (b) and encoding an amino acid sequence which is at least 70% identical to the amino acid sequence encoded by the DNA sequence of (a) or (b);
- [0019]** (d) DNA sequences, the nucleotide sequence of which is degenerated as a result of the genetic code to a nucleotide sequence of a DNA sequence as defined in any one of (a) to (c); and
- [0020]** (e) DNA sequences encoding a fragment of a protein encoded by a DNA sequence of any one of (a) to (d).

**[0021]** One embodiment of the invention is a protein with SEQ ID NO:2. Another embodiment of the invention is a protein with SEQ ID NO:4. A further embodiment of the invention concerns a protein with SEQ ID NO:6.

**[0022]** A further aspect of the invention is the use of the aforementioned proteins, or biologically active fragments thereof, to modulate and/or inhibit members of the TNF receptor superfamily such as CD40, CD30 or TNF-receptor II in their signaling activity and/or CD40-induced NF- $\kappa$ B activation and/or JUN-kinase activity.

**[0023]** The isolated functional protein according to the invention and/or a functional fragment thereof can be used to treat TRAF-related, CD40-related, NF- $\kappa$ B related and/or Jun (kinase)

related diseases. Such diseases include atherosclerosis, arthritis, multiple sclerosis, systemic lupus erythematosus (“SLE”) and/or graft rejection.

[0024] In addition, the isolated functional protein according to the invention and/or a functional fragment thereof can be used to sensitize tumor cells to anti-tumor treatments.

[0025] Another aspect of the invention is the use of aforementioned proteins or biologically active fragments thereof to screen for compounds that interfere in the interactions of the proteins or fragments with other protein components of the CD40, TRAF or NF- $\kappa$ B related signaling pathway.

[0026] The invention also relates to a method for identifying and obtaining an activator or inhibitor of CD40 receptor associated proteins comprising the steps of:

[0027] (a) combining a compound to be screened with a reaction mixture containing the protein of the invention and a readout system capable of interacting with the protein under suitable conditions;

[0028] (b) maintaining the reaction mixture in the presence of the compound or a sample comprising a plurality of compounds under conditions which permit interaction of the protein with the read out system; and

[0029] (c) identifying or verifying a sample and compound, respectively, which leads to suppression or activation of the read out system.

[0030] As used herein, the term “read out system” means a DNA sequence which upon transcription and/or expression in a cell, tissue or organism provides for a scorable and/or selectable phenotype. Such read out systems are well known to those skilled in the art and comprise, for example, recombinant DNA molecules and marker genes as previously described herein.

[0031] The term “plurality of compounds” in a method of the invention is to be understood as a plurality of substances which may or may not be identical.

[0032] The compound or plurality of compounds may be comprised in, for example, samples, *e.g.*, cell extracts from animals or microorganisms. Furthermore, the compound(s) may be known in the art but hitherto not known to be capable of suppressing or activating CD-40 receptor associated interacting proteins. The reaction mixture may be a cell free extract or may comprise a cell or tissue culture. Suitable set ups for the method of the invention are known to the person skilled in the art and are, for example, generally described in Alberts *et al.*, Molecular Biology of the

Cell, (3<sup>rd</sup> ed. 1994). The plurality of compounds may be, *e.g.*, added to the reaction mixture, culture medium or injected into the cell.

**[0033]** If a sample containing a compound or a plurality of compounds is identified in the method of the invention, then it is either possible to isolate the compound from the original sample identified as containing the compound capable of suppressing or activating CD40 receptor associated proteins, or one can further subdivide the original sample, for example, if it consists of a plurality of different compounds, so as to reduce the number of different substances per sample and repeat the method with the subdivisions of the original sample. Depending on the complexity of the samples, the steps described above can be performed several times, preferably until the sample identified according to the method of the invention only comprises a limited number of or only one substance(s). Preferably, the sample comprises substances of similar chemical and/or physical properties, and most preferably the substances are identical. The compounds which can be tested and identified according to a method of the invention may be expression libraries, *e.g.*, cDNA expression libraries, peptides, proteins, nucleic acids, antibodies, small organic compounds, hormones, peptidomimetics, PNAs or the like (Milner, *Nature Medicine* 1 (1995), 879-880; Hupp, *Cell* 83 (1995), 237-245; Gibbs, *Cell* 79 (1994), 193-198 and references cited *supra*).

**[0034]** Another aspect of the invention involves DNA molecules, also called nucleic acid sequences, encoding for the aforementioned proteins, preferably a nucleic acid sequence with 70-100% homology to the DNA sequence depicted in SEQ ID NO:1 and/or a nucleic acid sequence with 70-100% homology to the DNA sequence depicted in SEQ ID NO:3 or, in the alternative, a nucleic acid sequence with 70-100% homology to the DNA sequence depicted in SEQ ID NO:5.

**[0035]** "Homology," in this context, means that the respective nucleic acid molecules or encoded proteins are functionally and/or structurally equivalent. The nucleic acid molecules that are homologous to the nucleic acid molecules described above and that are derivatives of the nucleic acid molecules are, for example, variations of the nucleic acid molecules which represent modifications having the same biological function, in particular encoding proteins with the same or substantially the same biological function. They may be naturally occurring variations, such as sequences from other varieties or species, or mutations. These mutations may occur naturally or may

be obtained by mutagenesis techniques. The allelic variations may be naturally occurring allelic variants as well as synthetically produced or genetically engineered variants.

[0036] The proteins encoded by the various derivatives and variants of the above-described nucleic acid molecules have similar common characteristics, such as biological activity, molecular weight, immunological reactivity, conformation, etc., as well as physical properties, such as electrophoretic mobility, chromatographic behavior, sedimentation coefficients, pH optimum, temperature optimum, stability, solubility, spectroscopic properties, etc.

[0037] A further embodiment of the invention concerns a method for identifying and obtaining CD40 receptor associated proteins comprising the LexA two-hybrid system wherein LexA DNA-binding domain as a bait and a HeLa cell fusion library in plasmid pJG45 as prey is used. A DNA sequence encoding a CD40 receptor associated protein obtainable by the method belongs to the invention as well.

[0038] The present invention also relates to vectors, particularly plasmids, cosmids, viruses, bacteriophages and other vectors used conventionally in genetic engineering that contain a nucleic acid molecule according to the invention. Methods which are well known to those skilled in the art can be used to construct various plasmids and vectors; see, for example, the techniques described in Sambrook, *Molecular Cloning A Laboratory Manual*, Cold Spring Harbor Laboratory (1989) N.Y.

[0039] Alternatively, the nucleic acid molecules and vectors of the invention can be reconstituted into liposomes for delivery to target cells.

[0040] In a preferred embodiment, the nucleic acid molecule present in the vector is operably linked to a control sequence or control sequences that allow for the expression of the nucleic acid molecule in prokaryotic and/or eukaryotic cells.

[0041] The term "control sequence" refers to regulatory DNA sequences which are necessary to affect the expression of coding sequences to which they are ligated. The nature of such control sequences differs depending upon the host organism. In procaryotes, control sequences generally include promoter, ribosomal binding site, and terminators. In eucaryotes generally, control sequences include promoters, terminators and, in some instances, enhancers, transactivators or transcription factors. The term "control sequence" is intended to include, at a minimum, all



components the presence of which are necessary for expression, and may also include additional advantageous components.

[0042] The term “operably linked” refers to a juxtaposition wherein the components so described are in a relationship permitting them to function in their intended manner. A control sequence “operably linked” to a coding sequence is ligated in such a way that expression of the coding sequence is achieved under conditions compatible with the control sequences. If the control sequence is a promoter, a double-stranded nucleic acid is used.

[0043] Thus, the vector of the invention is preferably an expression vector. An “expression vector” is a construct that can be used to transform a selected host cell and provides for expression of a coding sequence in the selected host. Expression vectors can, for instance, be cloning vectors, binary vectors or integrating vectors. Expression comprises transcription of the nucleic acid molecule preferably into a translatable mRNA. Regulatory elements ensuring expression in prokaryotic and/or eukaryotic cells are well known to those skilled in the art.

[0044] The present invention furthermore relates to host cells comprising a vector as described above or a nucleic acid molecule according to the invention wherein the nucleic acid molecule is foreign to the host cell.

[0045] By “foreign” it is meant that the nucleic acid molecule is either heterologous with respect to the host cell, this means derived from a cell or organism with a different genomic background, or is homologous with respect to the host cell but located in a different genomic environment than the naturally occurring counterpart of the nucleic acid molecule. This means that, if the nucleic acid molecule is homologous with respect to the host cell, it is not located in its natural location in the genome of the host cell, in particular it is surrounded by different genes. In this case, the nucleic acid molecule may be either under the control of its own promoter or under the control of a heterologous promoter. The vector or nucleic acid molecule according to the invention which is present in the host cell may either be integrated into the genome of the host cell or it may be maintained in some form extra-chromasomally. In this respect, it is also to be understood that the nucleic acid molecule of the invention can be used to restore or create a mutant gene via homologous recombination (Paszkowski (ed.), *Homologous Recombination and Gene Silencing in Plants*, Kluwer Academic Publishers (1994)).

[0046] The host cell can be any prokaryotic or eukaryotic cell, such as bacterial, insect, fungal, plant or animal cells. Preferred fungal cells are, for example, those of the genus *Saccharomyces*, in particular those of the species *S. cerevisiae*.

[0047] The invention also includes a method for preparing CD40 receptor associated proteins which method comprises the cultivation of host cells according to the invention which, due to the presence of a vector or a nucleic acid molecule according to the invention, are able to express such a protein, under conditions which allow expression of the protein and recovering of the so-produced protein from the culture.

[0048] The present invention furthermore relates to proteins encoded by the nucleic acid molecules according to the invention or produced or obtained by the herein described methods, and to functional and/or immunologically active fragments of such CD40 receptor associated proteins. The proteins and polypeptides of the present invention are not necessarily translated from a designated nucleic acid sequence; the polypeptides may be generated in any manner, including for example, chemical synthesis, or expression of a recombinant expression system, or isolation from a suitable viral system. The polypeptides may include one or more analogs of amino acids, phosphorylated amino acids or unnatural amino acids. Methods of inserting analogs of amino acids into a sequence are known in the art. The polypeptides may also include one or more labels, which are known to those skilled in the art. In this context, it is also understood that the proteins according to the invention may be further modified by conventional methods known in the art. By providing the proteins according to the present invention it is also possible to determine fragments which retain biological activity, namely the mature, processed form. This allows the construction of chimeric proteins and peptides comprising an amino sequence derived from the protein of the invention, which is crucial for its binding activity. The other functional amino acid sequences may be either physically linked by, *e.g.*, chemical means to the proteins of the invention or may be fused by recombinant DNA techniques well known in the art.

[0049] Furthermore, the present invention relates to antibodies specifically recognizing a CD40 receptor associated protein according to the invention or parts, *i.e.*, specific fragments or epitopes of such a protein. The antibodies of the invention can be used to identify and isolate other CD40 receptor associated proteins and genes in any organism. These antibodies can be monoclonal

antibodies, polyclonal antibodies or synthetic antibodies as well as fragments of antibodies, such as Fab, Fv or scFv fragments, etc. Monoclonal antibodies can be prepared, for example, by the techniques as originally described in Köhler and Milstein, *Nature* 256 (1975), 495, and Galfré, *Meth. Enzymol.* 73 (1981), 3, which comprise the fusion of mouse myeloma cells to spleen cells derived from immunized mammals. Furthermore, antibodies or fragments thereof to the aforementioned peptides can be obtained by using methods which are described, *e.g.*, in Harlow and Lane "Antibodies, A Laboratory Manual," CSH Press, Cold Spring Harbor, 1988. These antibodies can be used, for example, for the immunoprecipitation and immunolocalization of proteins according to the invention as well as for the monitoring of the synthesis of such proteins, for example, in recombinant organisms and for the identification of compounds interacting with the protein according to the invention. For example, surface plasmon resonance as employed in the BIAcore system can be used to increase the efficiency of phage antibodies selections, yielding a high increment of affinity from a single library of phage antibodies which bind to an epitope of the protein of the invention (Schier, *Human Antibodies Hybridomas* 7 (1996), 97-105; Malmberg, *J. Immunol. Methods* 183 (1995), 7-13). In many cases, the binding phenomena of antibodies to antigens is equivalent to other ligand/anti-ligand binding.

**[0050]** The invention also relates to a diagnostic composition comprising at least one of the aforementioned nucleic acid molecules, vectors, proteins, antibodies or compounds and optionally suitable means for detection.

**[0051]** The diagnostic compositions may be used for methods for detecting expression of related CD40 receptor associated proteins by detecting the presence of the corresponding mRNA which comprises isolation of mRNA from a cell and contacting the mRNA so obtained with a probe comprising a nucleic acid probe as described above under hybridizing conditions, detecting the presence of mRNA hybridized to the probe, and thereby detecting the expression of the protein in the cell. Further methods of detecting the presence of a protein according to the present invention comprises immuno-techniques well known in the art, for example, enzyme linked immunosorbent assay.

**[0052]** The invention also relates to a pharmaceutical composition comprising one or more of the above mentioned proteins or fragments in a biologically active amount for the treatment of

CD40, TRAF and/or NF- $\kappa$ B related diseases such as atherosclerosis, arthritis, multiple sclerosis, systemic lupus erythematosus, graft rejection and the like.

[0053] In another aspect, the present invention relates to a pharmaceutical composition comprising one or more compounds obtainable by the above-mentioned screening method for the treatment of CD40, TRAF and/or NF- $\kappa$ B related diseases such as atherosclerosis, arthritis, multiple sclerosis, systemic lupus erythematosus, graft rejection and the like.

#### BRIEF DESCRIPTION OF THE FIGURES

[0054] FIG. 1: Schematic representation of TTRAP (=CD40 receptor associated protein) and the deletion mutants of TTRAP used in two-hybrid assays. The deletion mutants consist of the following amino acids of the original TTRAP sequence: 54 to 362 (4F2), 54 to 273 (4F2d3), 54 to 236 (4F2d2) and 54 to 140 (4F2d1). (CaM= calcium calmodulin binding region.)

[0055] FIG. 2: Northern blot analysis of (a) human tissue, using a human TTRAP probe; (b) adult mouse tissue, using a mouse TTRAP probe; (c) embryonic mouse tissue, using a mouse TTRAP probe. The hybridization of GAPDH is used as a control.

#### DETAILED DESCRIPTION OF THE INVENTION

##### DEFINITIONS

[0056] The following definitions are provided in order to further illustrate and define the meaning and scope of the various terms used in the current description.

[0057] As used herein, "homology to TRAF-proteins" means that the typical structural features found in the current TRAF proteins (TRAF1 - TRAF6) are present. These features comprise a RING finger motif at the amino terminus followed by five or more zinc fingers and a so-called TRAF domain known to a person skilled in the art.

[0058] As used herein, "CD40," it is not to limit the scope to CD40 only, but also includes other receptors of the TNF receptor superfamily such as CD30 or TNF receptor II, unless indicated to the contrary.

[0059] The term "treatment" or "treating" or "treat" means any treatment of a disease in a mammal, including: (1) preventing the disease causing the clinical symptoms of the disease not to

develop; (2) inhibiting the disease arresting the development of the clinical symptoms; and/or (3) relieving the disease causing the regression of clinical symptoms.

**[0060]** The term “effective amount” means a dosage sufficient to provide treatment for the disease state being treated. This will vary depending on the patient, the disease and the treatment being effected.

**[0061]** “Capable of interacting” means that a protein can form a complex with another protein, as can be measured using a yeast two-hybrid system, or with co-immunoprecipitation, or with equivalent systems known to people skilled in the art.

**[0062]** “Functional” protein or fragment means a protein or fragment that is capable to interact with the cytoplasmic part of CD40, or with another protein of the CD40 and/or NF- $\kappa$ B related pathway.

**[0063]** The “cytoplasmic part of CD40” means a part comprising the 62 carboxy terminal amino acids of human CD40 (amino acid 216-277; Stamenkovic *et al.* 1989), or the homologous mouse sequence, or another homologous sequence with a similar biological activity.

**[0064]** The term “expression” means the production of a protein or nucleotide sequence in the cell. However, the term also includes expression of the protein in a cell-free system. It includes transcription into an RNA product, post-transcriptional modification and/or translation to a protein product or polypeptide from a DNA encoding that product, as well as possible post-translational modifications. Depending on the specific constructs and conditions used, the protein may be recovered from the cells, from the culture medium or from both. For the person skilled in the art it is well known that it is not only possible to express a native protein but also to express the protein as fusion polypeptides or to add signal sequences directing the protein to specific compartments of the host cell, *e.g.*, ensuring secretion of the peptide into the culture medium, etc. Furthermore, such a protein and fragments thereof can be chemically synthesized and/or modified according to standard methods.

**[0065]** The terms “protein” and “polypeptide,” as used herein, are interchangeable. “Polypeptide” refers to a polymer of amino acids (amino acid sequence) and does not refer to a specific length of the molecule. Thus, peptides and oligopeptides are included within the definition of polypeptide. This term also refers to or includes post-translational modifications of the

polypeptide, for example, glycosylations, acetylations, phosphorylations and the like. Included within the definition are, for example, polypeptides containing one or more analogs of an amino acid (including, for example, unnatural amino acids, etc.), polypeptides with substituted linkages, as well as other modifications known in the art, both naturally occurring and non-naturally occurring.

**[0066]** The terms “gene(s),” “polynucleotide,” “nucleic acid sequence,” “nucleotide sequence,” “DNA sequence” or “nucleic acid molecule(s),” as used herein, refer to a polymeric form of nucleotides of any length, either ribonucleotides or deoxyribonucleotides. This term refers only to the primary structure of the molecule. Thus, this term includes double- and single-stranded DNA and RNA. It also includes known types of modifications, for example, methylation, “caps” substitution of one or more of the naturally occurring nucleotides with an analog. Preferably, the DNA sequence of the invention comprises a coding sequence encoding the above defined CD40 receptor associated protein.

**[0067]** A “coding sequence” is a nucleotide sequence which is transcribed into mRNA and/or translated into a polypeptide when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a translation start codon at the 5'-terminus and a translation stop codon at the 3'-terminus. A coding sequence can include, but is not limited to, mRNA, cDNA, recombinant nucleotide sequences or genomic DNA, while introns may be present as well under certain circumstances.

**[0068]** The term “functional fragment of a sequence” or “functional part of a sequence” means a truncated sequence of the original sequence referred to. The truncated sequence (nucleic acid or protein sequence) can vary widely in length; the minimum size being a sequence of sufficient size to provide a sequence with at least a comparable function and/or activity of the original sequence referred to, while the maximum size is not critical. In some applications, the maximum size usually is not substantially greater than that required to provide the desired activity and/or function(s) of the original sequence. Typically, the truncated amino acid sequence will range from about 5 to about 60 amino acids in length. More typically, however, the sequence will be a maximum of about 50 amino acids in length, preferably a maximum of about 30 amino acids. It is usually desirable to select sequences of at least about 10, 12 or 15 amino acids, up to a maximum of about 20 or 25 amino acids.

**[0069]** Furthermore, folding simulations and computer redesign of structural motifs of the protein of the invention can be performed using appropriate computer programs (Olszewski, *Proteins* 25 (1996), 286-299; Hoffman, *Comput. Appl. Biosci.* 11 (1995), 675-679). Computer modeling of protein folding can be used for the conformational and energetic analysis of detailed peptide and protein models (Monge, *J. Mol. Biol.* 247 (1995), 995-1012; Renouf, *Adv. Exp. Med. Biol.* 376 (1995), 37-45). In particular, the appropriate programs can be used for the identification of interactive sites of the inventive protein, its receptor, its ligand or other interacting proteins by computer assisted searches for complementary peptide sequences (Fassina, *Immunomethods* 5 (1994), 114-120). Further, appropriate computer systems for the design of protein and peptides are described in the prior art, for example in Berry, *Biochem. Soc. Trans.* 22 (1994), 1033-1036; Wodak, *Ann. N. Y. Acad. Sci.* 501 (1987), 1-13; and Pabo, *Biochemistry* 25 (1986), 5987-5991. The results obtained from the above-described computer analysis can be used for, *e.g.*, the preparation of peptidomimetics of the protein of the invention or fragments thereof. Such pseudopeptide analogues of the natural amino acid sequence of the protein may very efficiently mimic the parent protein (Benkirane, *J. Biol. Chem.* 271 (1996), 33218-33224). For example, incorporation of easily available achiral  $\alpha$ -amino acid residues into a protein of the invention or a fragment thereof results in the substitution of amide bonds by polymethylene units of an aliphatic chain, thereby providing a convenient strategy for constructing a peptidomimetic (Banerjee, *Biopolymers* 39 (1996), 769-777). Superactive peptidomimetic analogues of small peptide hormones in other systems are described in the prior art (Zhang, *Biochem. Biophys. Res. Commun.* 224 (1996), 327-331). Appropriate peptidomimetics of the protein of the present invention can also be identified by the synthesis of peptidomimetic combinatorial libraries through successive amide alkylation and testing the resulting compounds, *e.g.*, for their binding and immunological properties. Methods for the generation and use of peptidomimetic combinatorial libraries are described in the prior art, for example, in Ostresh, *Methods in Enzymology* 267 (1996), 220-234 and Dorner, *Bioorg. Med. Chem.* 4 (1996), 709-715.

**[0070]** Furthermore, a three-dimensional and/or crystallographic structure of the protein of the invention can be used for the design of peptidomimetic inhibitors of the biological activity of the protein of the invention (Rose, *Biochemistry* 35 (1996), 12933-12944; Rutenber, *Bioorg. Med. Chem.* 4 (1996), 1545-1558).

[0071] “Compound” means any chemical or biological compound, including simple or complex inorganic or organic molecules, peptides, peptido-mimetics, proteins, antibodies, carbohydrates or nucleic acids, that interferes with the interaction of a protein depicted in SEQ ID NO:2, SEQ ID NO:4, or SEQ ID NO:6 with a compound of the CD40 and/or NF- $\kappa$ B related pathway.

[0072] As used herein, the term “composition” refers to any composition such as a pharmaceutical composition comprising as an active ingredient a functional protein according to the present invention possibly in the presence of suitable excipients known to the skilled man and may thus be administered in the form of any suitable composition as detailed below by any suitable method of administration within the knowledge of a skilled man. The preferred route of administration is parenterally. In parenteral administration, the compositions of this invention will be formulated in a unit dosage injectable form such as a solution, suspension or emulsion, in association with a pharmaceutically acceptable excipient. Such excipients are inherently nontoxic and non-therapeutic. Examples of such excipients are saline, Ringer’s solution, dextrose solution and Hank’s solution. Non-aqueous excipients such as fixed oils and ethyl oleate may also be used. A preferred excipient is 5% dextrose in saline. The excipient may contain minor amounts of additives such as substances that enhance isotonicity and chemical stability, including buffers and preservatives.

[0073] The functional protein of the invention is administered at a concentration that is therapeutically effective to prevent allograft rejection, graft versus host disease (“GVHD”), allergy and autoimmune diseases. The dosage and mode of administration will depend on the individual. Generally, the compositions are administered so that the functional protein is given at a dose between 1 mg/kg and 10 mg/kg, more preferably between 10 mg/kg and 5 mg/kg, and most preferably between 0.1 and 2 mg/kg. Preferably, it is given as a bolus dose. Continuous short time infusion (during 30 minutes) may also be used. The compositions comprising the functional protein according to the invention may be infused at a dose between 5 and 20 mg/kg/minute, more preferably between 7 and 15 mg/kg/minute.

[0074] According to a specific case, the “therapeutically effective amount” of the functional protein according to the invention needed should be determined as being the amount



sufficient to cure the patient in need of treatment or at least to partially arrest the disease and its complications. Amounts effective for such use will depend on the severity of the disease and the general state of the patient's health. Single or multiple administrations may be required depending on the dosage and frequency as required and tolerated by the patient.

**[0075]** With regard to the use of the functional protein of the present invention to prevent allograft rejection, it should be stressed that the proteins of the present invention or the compositions comprising the same may be administered before, during or after the organ transplantation as is desired from case to case. In case the protein or the compositions comprising the same are administered directly to the host, treatment will preferably start at the time of the transplantation and continue afterwards in order to prevent the activation and differentiation of host T cells against the major histocompatibility complex ("MHC") on the allograft. In case the donor organ is *ex vivo* perfused with the functional protein according to the invention or the compositions comprising the same, treatment of the donor organ *ex vivo* will start before the time of the transplantation of the donor organ in order to prevent the activation and differentiation of host T cells against the MHC on the allograft

**[0076]** The invention is further explained by way of the following illustrative examples:

## EXAMPLES

### Example 1: isolation of the CD40 Interacting Proteins

#### Yeast two-hybrid screening.

**[0077]** The two-hybrid assay was performed by the interaction trap cloning method, which is often referred to as the LexA two-hybrid system (Gyuris *et al.*, 1993). The DNA encoding the cytoplasmic part of CD40 (62 amino acids, from residue 216 to 277, where the open reading frame ends, according to the sequence and numbering as given in Stamenkovic *et al.* (1989)) was generated by PCR and inserted into the EcoRI-SalI digested pEG202 vector (Gyuris *et al.*, 1993), in frame with the LeXA DNA-binding domain (hereinafter the "bait plasmid"). Screening was performed using a HeLa cell fusion library in the plasmid pJG45 (hereinafter the "prey plasmid") obtained from the laboratory of R. Brent (Harvard Med. School, Boston, MA, USA). Transformation of EGY48 yeast (MAT alpha, *his3*, *trp1*, *ura3-52*, *leu2::pLeu2-LexAop6*) with the prey plasmid, the bait plasmid and

the p8op-LacZ (Clontech) reporter plasmid was performed by the Lithium Acetate transformation method (Gietz *et al.*, 1995). The two-hybrid screening was conducted as described in the manual distributed by the laboratory of R. Brent (published in “Gene probes-A practical approach, Oxford University press”).

#### Results of the two-hybrid screening.

**[0078]** Yeast containing bait plasmid and lacZ reporter plasmid was transformed with 20 microgram prey library plasmid and plated on glucose medium lacking tryptophan, histidine and uracil to select for the presence of all three plasmids. In total, approximately  $1.5 \times 10^6$  colonies were obtained. The colonies were harvested and frozen at  $-70^{\circ}\text{C}$  in a glycerol solution (65% v/v glycerol; 0.1 M  $\text{MgSO}_4$ , 25 mM Tris pH 7.4). From these stocks,  $20 \times 10^6$  colony forming units were plated on galactose medium lacking leucine, tryptophan, histidine and uracil to screen for protein-protein interaction. Yeast colonies growing on the latter selective medium were further checked for interaction by screening for blue/white staining on medium containing X-gal and galactose. The colonies displaying the following phenotype were picked for further analysis: i) no growth on glucose containing medium which lacks leucine, ii) growth on galactose containing medium lacking leucine, iii) white on medium containing glucose and X-gal, and iv) blue on medium containing galactose and X-gal.

**[0079]** Plasmids were isolated from the yeast with the proper phenotype. Analysis of the obtained prey plasmids revealed that the entire screening had finally resulted in the isolation of three different cDNA inserts. Sequencing of the clones showed, in addition to a partial cDNA for TRAF3, two novel cDNA's, termed TTRAP and 4C4.

#### Isolation of the full-length cDNA

**[0080]** Full-length human TTRAP cDNA was obtained by screening a HUVEC cDNA library with the TTRAP fragment as a probe. A cDNA of about 2 kb was isolated, with an open reading frame of 1086 nucleotides encoding for a protein of 362 amino acids (SEQ ID NO:2).

[0081] The mouse TTRAP homologue was obtained by screening the EST database and aligning the homologous sequences. Human and mouse TTRAP are approximately 65% identical and 70% similar on the amino acid level. The mouse sequence is shown in SEQ ID NO:3.

#### Example 2: Sequence Analysis of the cDNA's

[0082] Nucleotide sequence analysis was carried out using dye terminator mix and a 310 Genetic analyzer from Perkin Elmer. The nucleotide sequence of TTRAP is shown in SEQ ID NO:1 whereas the sequence of 4C4 is shown in SEQ ID NO:5.

[0083] The TTRAP sequence shows a low homology (30% similarity at amino acid level) with Nocturnin, a protein that is expressed in the photoreceptor of the eye of *Xenopus laevis* (Green and Beshare, 1996). The partial sequence of the mouse homologue of Nocturnin is known (Puech *et al.*, 1997). Additionally, there is some homology with EST sequences (e.g., Genbank EST c23016, aa162513, aa571061, t87026, h45114, aa196281, h94108 and aa337396) and with the C-terminal part of the yeast transcription factor *CCR4* (Malvar *et al.*, 1992). Although these homologies are low, it is not excluded that a human homologue of these proteins would bind to the cytoplasmic domain of CD40.

[0084] It is interesting to note that, unexpectedly, TTRAP neither 4C4 show any significant homology with TRAF's or other proteins known to interact with CD40.

#### Example 3: Study of the Interaction of TTRAP Protein, 4c4 Protein and TTRAP Protein Fragments with Other Proteins Using a Yeast Two-hybrid Interaction Assay

[0085] The potential binding of TTRAP to other proteins was assessed using a yeast two-hybrid assay. The experimental outline is similar to the one described for the two-hybrid screening. However, here the plasmids for bait, prey and lacZ reporter were transformed simultaneously into the EGY48 yeast strain. Positive interaction was determined either by the growth phenotype (growth on medium lacking leucine in the presence of galactose, and not in the presence of glucose) or by the blue/white staining on X-gal containing plates (blue colonies only on galactose containing plates, not on glucose containing plates). cDNA's for TRAF2 and for the cytoplasmic regions of CD30, CD40 and TNF-RII were generated by PCR using the pfu polymerase

(Promega). PCR fragments encoding RIP, TRADD and FADD were cloned in pCDNA3 (Invitrogen, Carlsbad, CA). cDNA of TRAF3 was obtained from the laboratory of Dixit, Dept Pathol., Univ. Michigan Med. School, MI, USA. The color formation was evaluated as strong and fast (++), strong but slow (+), weak and slow (+/-), none (-) or not determined (nd).

[0086] The results for TTRAP protein and TTRAP fragments used are summarized in Table I and FIG. 1, respectively.

[0087] Table I

	<b>TTRAP</b>	<b>4F2</b>	<b>4F2d3</b>	<b>4F2d2</b>	<b>4F2d1</b>	<b>4C4</b>	<b>---</b>
<b>CD40</b>	++	++	+/-	+/-	+/-	+	-
<b>CD30</b>	++	++	+/-	+/-	+/-	+	-
<b>TNF-RII</b>	+	+	-	-	-	+	-
<b>LMP-1</b>	-	-	nd	nd	nd	-	-
<b>TRAF2</b>	+	+	nd	nd	nd	nd	-
<b>TRAF3</b>	+	+	+/-	+/-	+/-	nd	-
<b>RIP</b>	++	++	+/-	+/-	+/-	nd	+/-
<b>TRADD</b>	+	nd	nd	nd	nd	nd	-
<b>FADD</b>	-	nd	nd	nd	nd	nd	-
<b>4F2</b>	++	++	-	-	-	+	-
<b>4C4</b>	++	++	-	-	-	+	-

[0088] TTRAP, as well as the longest TTRAP fragment (aa 54 - 362), shows a strong interaction with CD40, CD30, RIP and with 4C4 and a weaker interaction with TNF-RII and TRAF3. Remarkably, TTRAP can also self-associate. TTRAP fragments, missing the C-terminal end (aa 274 - 362), show only a weak interaction.

[0089] Yeast transformed with TTRAP or 4F2 as a prey and TRAF2 as a bait gives a positive interaction phenotype similar to the one observed for TTRAP and TRAF3. This clearly indicates that TTRAP/4F2 also associates with TRAF2.

[0090] Because of the interaction of TTRAP with TRAF2 and 3 in two-hybrid assays in yeast, and with TRAF5 in co-immunoprecipitations from mammalian cells (see example 6) the conclusion is justified that TTRAP binds all other TRAFs as well. Deletion mutants of TRAFs are

constructed to find out whether the region commonly denominated as the TRAF domain is responsible for the binding of the TRAFs to TTRAP. In a similar approach, deletion mutants of TTRAP are constructed to find out which region in the protein is required and sufficient for the interaction with TRAFs.

[0091] 4C4 protein is interacting with CD40, CD30, TNF-RII, with the longest fragment of TTRAP and with a deletion mutant of TRAF3 which still contains the largest part of the TRAF domain (from aa 380 to the carboxy terminal end of the protein). A smaller form of 4C4 (from amino acid 2 - amino acid 245 in SEQ ID NO:6) is also capable to interact with CD40.

#### Example 4: Expression Pattern of TTRAP and 4C4

[0092] The TTRAP gene is widely expressed, as was already indicated by the presence of several partial TTRAP cDNA's in the EST sequence data base. The TTRAP expression was analyzed by Northern blot analysis against mRNA from different tissues, both from human and mouse (FIG. 2). Human TTRAP is present as a 2.2 kb transcript in all tissues tested. In addition to the 2.2 kb transcript, an additional 1.7 kb transcript is present in a testis sample (FIG. 2A).

[0093] Human TTRAP expression was further tested and found in the B-cell lines BJAB (Menezes *et al.*, 1975) and DG75 (Ben-Bassat *et al.*, 1977), in the Jurkat T-cell line and in HUVECs (Human umbilical vein endothelial cell).

[0094] For mouse TTRAP, two transcripts, one of 2.2 kb and one of 3.8 kb, were found on a murine multiple Northern blot (FIG. 2B). Mouse TTRAP mRNA was also detected in all tissues tested, be it to a lower extent in skeletal muscle. Both mouse transcripts are not only present in adult animals, but can also be detected in mouse embryo's from 7 to 17 days post coitus. These results are an indication that TTRAP plays an important role in early development.

[0095] On a multiple tissue Northern blot, a 4C4 probe recognizes 3 transcripts of 1.6 kb, 3.5 kb and 7.5kb. All three mRNA's are present in spleen, thymus, prostate, testis, ovary, small intestine, colon and peripheral blood leukocytes. The expression of the 3.5 kb transcript is most prominent in testis. In ovary, the signal of the 7.5 kb mRNA is strongest.

#### Example 5: Co-immunoprecipitation of TTRAP with TNF-RII

[0096] It has been shown by two-hybrid analysis in yeast that TTRAP interacts with different receptors of the TNF-receptor superfamily, i.e., CD30, CD40 and TNF-RII (see, Table I). To confirm these results for one of these receptors in mammalian cells, co-immunoprecipitation of TTRAP with TNF-RII was carried out. In a typical experiment, 293T cells were transfected with HA-tagged TTRAP and TNF-RII (both in pCDNA expression vectors). TNF-RII was immunoprecipitated with the utr4 monoclonal antibody (gift from Roche, Basel). By Western blot on the immunoprecipitated fraction it was confirmed that TTRAP interacts with TNF-RII.

[0097] The method used hereto is as follows: A 9cm dish with 293T Human Embryonic Kidney cells (approx. 30-50% confluent) is transfected with approximately 2 microgram of each expression plasmid. 24-48 hours after transfection, the cells are harvested and lysed in 300 microliter lysis buffer (50 mM Tris/HCl pH 7.4, 200 mM NaCl, 10% glycerol, 0.2% NP-40, 50 mM NaF, 1 mM  $\text{Na}_3\text{P}_2\text{O}_7$ , 1 mM  $\text{Na}_3\text{VO}_4$  and protease inhibitors). The cell suspension is incubated for 20 minutes on ice. Cellular debris is pelleted for 10 minutes in an Eppendorf centrifuge, at 14.000 rpm and 4°C, and the supernatant is transferred into a fresh tube. 5 microgram antibody is added to the lysate, and incubated for 3 hours at 4°C, on a mixing platform. The samples are supplemented with 20 microliter of a 50% slurry of protein A or G beads and the incubation is continued for 1 hour. The beads are pelleted for 1 minute at minimal speed (approx. 500 rpm) and the supernatant is removed. The beads are resuspended in 750 microliter lysis buffer and immediately centrifuged again. This washing procedure is repeated, but this time samples are incubated for 10-15 minutes at 4°C on a mixing platform. This is repeated twice more, to come to a total of 4 buffer changes. After the last wash, as much buffer as possible is to be removed, without sucking up Sepharose beads. After addition of 20 microliter sample buffer, the samples are stored at -20°C or processed for SDS/PAGE. 10 microliters of each sample is analyzed to determine the protein CoIP and 3 microliters to verify the IP.

[0098] Interestingly, it was found that the association of TTRAP with the TNF receptor in an immunoprecipitation protocol from mammalian cells is preserved only if the immunoprecipitation procedure was performed in the presence of inhibitors of protein phosphatases, i.e., 50 mM NaF, 1

mM  $\text{Na}_3\text{P}_2\text{O}_7$ , 1 mM  $\text{Na}_3\text{VO}_4$ . This suggests that TTRAP and/or the receptor are phosphorylated, and that this phosphorylation is required for the interaction of both proteins.

#### Example 6: Interaction of TTRAP with TRAF3 and TRAF5

**[0099]** TTRAP was initially isolated in a two-hybrid screen in yeast, using the cytoplasmic domain of CD40 as bait. The interaction of TTRAP with CD40 could subsequently be confirmed by co-immunoprecipitation of both proteins, after over-expression in 293T human embryonic kidney cells. Because CD40 is known to associate with TRAFs in mammalian cells, it was investigated whether TTRAP could bind to TRAF3 and TRAF5.

**[00100]** The interaction of TTRAP with TRAF3 is disclosed in Table I. To investigate whether TTRAP could interact with TRAF5, a co-immunoprecipitation experiment was performed.

**[00101]** In this experiment, immunoprecipitation was carried out for Flag-tagged human TRAF5 (gift from Jun-ichiro Inoue, Dept Oncology, Inst. Med Sci., Univ. Tokyo, Japan) cells co-transfected with TRAF5 and TTRAP. Co-immunoprecipitation of HA-tagged TTRAP was detected by analyzing the immunoprecipitated proteins on Western blots. These results clearly demonstrate that TTRAP forms a complex with TRAF5 in mammalian cells.

#### Example 7: Interaction of TTRAP with CD40

**[00102]** To delineate the region in the cytoplasmic tail of CD40 that is important for TTRAP binding, deletion mutants of the receptor are made and are tested in co-immunoprecipitations and yeast two-hybrid assays. In a similar approach, deletion mutants of TTRAP are made to find out which region in the protein is necessary and sufficient for the interaction with CD40.

#### Example 8: Genomic TTRAP Sequence

**[00103]** Recently, the human genomic sequence for TTRAP has been deposited to the EMBL database (accession number HS30M3). The chromosome location of the genomic clone was mapped to 6p22.1-22.3. The TTRAP sequence within this clone was identified as a novel protein, termed dJ30M3.3 (acc number, CAA21141). The mRNA sequence starts at nucleotide 47,151 and ends at position 64,053. The protein encoding region starts at 47,168 and stops at 63,242,

comprising 7 exons. The translational start and stop sites match those that were determined for TTRAP by cDNA library screening and the amino acid sequence is identical to the human sequence as disclosed herein.

**[00104]** The human TTRAP sequence also aligns with the *C. elegans* gene product Y63D3A.4 (accession number CEY63D3A\_4). At the protein level, the sequences are approximately 30% identical. On the basis of the protein alignment it is assumed that this *C. elegans* protein is the nematode homologue of human TTRAP.

#### Example 9: Identification of Novel CD40-interacting Proteins

**[00105]** A two-hybrid screening in yeast was performed with the cytoplasmic tail of human CD40 as a bait and a HeLa cDNA library as a prey. In this screening, the proteins encoded by the cDNAs listed below, were picked up as a result of their interaction with CD40.

**[00106]** 1. A partial cDNA encoding human Bloom's syndrome protein (BLM) (accession number U39817). The complete cDNA sequence of BLM is 4,437 nucleotides. The protein encoding sequence starts at nucleotide 75 and stops at nucleotide 4,328. The size of the protein is 1,418 amino acids. The partial cDNA that was picked up in the two-hybrid screen with CD40 starts at nucleotide 529, which corresponds to amino acid 151. The 3'-end of this partial cDNA clone has not been determined yet.

**[00107]** 2. A partial cDNA encoding for human nuclear autoantigen (SP-100) (accession number M60618). The complete cDNA sequence of SP-100 is 1,879 nucleotides. The protein encoding sequence starts at nucleotide 32 and stops at nucleotide 1,474. The size of the protein is 481 amino acids. The partial cDNA that was picked up in the two-hybrid screen with CD40 starts at nucleotide 699, which corresponds to amino acid 223, and reaches till the end of the SP-100 cDNA sequence.

**[00108]** 3. A partial cDNA sequence for the human homologue of the mouse BP75 protein (accession number AF084259). The complete cDNA sequence of the mouse protein BP75 is 2,361 nucleotides. The protein encoding sequence



starts at nucleotide 211 and stops at nucleotide 2,166. The size of the protein is 652 amino acids. The partial cDNA that was picked up in the two-hybrid screen with CD40 aligns with the mouse cDNA sequence starting at nucleotide 1,070, which corresponds to amino acid 286, and goes to the end of the BP75 cDNA sequence. The mouse and the human cDNA sequences are 83% identical. The human genomic sequence of BP75 has also been deposited to the database (accession number Z99496). The cDNA of mouse BP75 is approximately 83% identical to the human genomic PAC clone (from nucleotide 122564-120820). The chromosome location of the human PAC clone was mapped to 6q22.1-22.33.

**Example 10: Isolation of DAXX as a CD40-interacting Protein by Yeast two-hybrid**

**[00109]** In a yeast two-hybrid assay with the cytoplasmic tail of human CD40 as a bait, a partial cDNA was picked up encoding the human Fas-binding protein (DAXX) (accession number AF039136). The complete cDNA sequence of DAXX is 2,487 nucleotides. The protein encoding sequence starts at nucleotide 148 and stops at nucleotide 2,370. The size of the protein is 741 amino acids. The partial cDNA that was picked up in the two-hybrid screen with CD40 starts at nucleotide 1,500, which corresponds to amino acid 451, and goes to the end of the DAXX cDNA sequence.

**[00110]** Fas and CD40 are both members of the TNF-Receptor superfamily. DAXX was originally isolated as a Fas-binding protein in a yeast two-hybrid screen (Yang *et al.*, Cell, 89, 1067-76, 1997). The protein was shown to interact specifically with the death domain of Fas. It was reported to play a role in apoptosis via the activation of the Jun N-terminal kinase. The authors examined the binding of a partial clone of human DAXX (from amino acid 501 till the end) to the cytoplasmic tail of mouse CD40 and could not detect interaction. In addition, an *in vitro* interaction assay of full length DAXX with glutathione S-transferase-CD40 ("GST-CD40") also turned out to be negative. Therefore, the authors conclude that DAXX does not associate with CD40.

**[00111]** The discrepancy between some of the findings reported in Yang *et al.* and the observation according to the current invention that the cytoplasmic tail of CD40 does interact with DAXX in yeast two-hybrid could be due to the following reasons: i) In the two-hybrid interaction

assay that was performed, Yang *et al.* use a shorter partial clone of DAXX than the one picked up according to the present invention. In addition, interaction of a partial human DAXX with human CD40 is observed in the underlying invention, whereas Yang *et al.*, use a partial human DAXX and mouse CD40; and ii) The *in vitro* binding assay that Yang *et al.* use may not be sensitive enough to detect the interaction of full-length DAXX and CD40 protein. Moreover, it is unclear whether in this assay Yang *et al.* use both proteins from the same species.

## REFERENCES

- Banchereau, J.; Bazan, F.; Blanchard, D.; Briere, F.; Galizzi, J.P.; Van Kooten, C.; Liu, Y.J.; Rousset, F.; Saeland, S. (1994). The CD40 antigen and its ligand. *Annu. Rev. Immunol.* 12, 881.
- Ben-Bassat, H.; Goldblum, N.; Mitrani, S.; Goldblum, T.; Yoffey, J.M.; Cohen, M.M.; Bentwich, Z.; Ramot, B.; Klein, E.; Klein, G. (1977). Establishment in continuous culture of a new type of lymphocyte from a "Burkitt like" malignant lymphoma (line D.G.-75). *Int. J. Cancer* 19, 27.
- Chaudhuri, A.; Orme, S.; Eilan, S.; Cherayl, B.J. (1997). CD40 mediated signals inhibit the binding of TNF receptor associated factor 2 to the CD40 cytoplasmic domain. *J. Immunol.* 159, 4244.
- Durie, F.H.; Fava, R.A.; Foy, T.M.; Aruffo, A.; Ledbetter, J.A.; Noelle, R.J. (1993). Prevention of collagen-induced arthritis with an antibody to gp39, the ligand for CD40. *Science* 261, 1328.
- Francis, D.A.; Karras, J.G.; Ke, X.Y.; Sen R.; Rothstein, T.L. (1995). Induction of the transcription factors NF-kappa B, AP-1 and NF-AT during B cell stimulation through the CD40 receptor. *Int. Immunol.* 7, 151.
- Fries, K.M.; Sempowski, G.D.; Gaspari, A.A.; Blieden, R.J.; Looney, R.J.; Philips, R.P. (1995). CD40 expression by human fibroblasts. *Clin. Immunol. Immunopath.* 77, 42.
- Galy, A.; Spits, H. (1992). CD40 is functionally expressed on human thymic epithelial cells. *J. Immunol.* 149, 775.
- Gaspari, A.A.; Sempowski, G.D.; Chess, P.; Gish, J.; Phillips, R.P. (1996). Human epidermal keratinocytes are induced to secrete interleukin-6 and costimulate T lymphocyte proliferation by a CD40-dependent mechanism. *Eur. J. Immunol.* 26, 1371.
- Gerritse, K.; Laman, J.D.; Noelle, R.J.; Arufo, A.; Ledbetter, J.A.; Boersma, W.J.A.; Claassen, E. (1996). CD40-CD40 ligand interactions in experimental allergic encephalomyelitis and multiple sclerosis. *Proc. Natl. Acad. Sci. USA* 93, 2499.
- Gietz, R.D.; Schiestl, R.H.; Willems, A.R.; Woods, R.A. (1995). Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast* ,11, 355.
- Green, C.B.; Beshare, J.C. (1996). Identification of a novel vertebrate circadian clock-regulated gene encoding the protein nocturnin. *Proc. Natl. Acad. Sci. USA* 93, 14884.

- Griggs, N.; Agersborg, S.; Noelle, R.J.; Ledbetter, J.; Linsley, P.; Tung, K. (1996). The relative contribution of the CD28 and gp39 costimulatory pathways in the clonal expansion and pathogenic acquisition of self-reactive T cells. *J. Exp. Med.* 180, 801.
- Gyuris, J.; Golemis, E.; Chertkov, H.; Brent, R. (1993). Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. *Cell* 75, 791.
- Hollenbaugh, D.; Mischel-Petty, N.; Edwards, C.P.; Simon, J.C.; Denfeld, R.W.; Kienerm, P.A.; Aruffo, A. (1995). Expression of functional CD40 by vascular endothelial cells. *J. Exp. Med.* 182, 33.
- Ishida, T.; Tojo, T.; Aoli, T.; Kobayashi, N.; Ohishi, T.; Watanabe, T.; Yamamoto, T.; Inoue, J. (1996). TRAF5, a novel tumor necrosis factor receptor associated factor family protein, mediated CD40 signalling. *Proc. Natl. Acad. Sci. USA* 93, 9437.
- Karmann, K.; Min, W.; Fanslow, W.C.; Prober, J.S. (1996). Activation and homologous desensitization of human endothelial cells by CD40 ligand, tumor necrosis factor and interleukin 1. *J. Exp. Med.* 184, 173.
- Kashiwada, M.; Shirakata, Y.; Inoue, J.-I.; Nakano, H.; Okazaki, K.; Okumura, K.; Yamamoto, T.; Nagaoka, H.; Takemori, T. (1998). Tumor Necrosis Factor Receptor-associated Factor 6 (TRAF6) stimulates extracellular signal-regulated kinase (ERK) activity in CD40 signalling along a Ras-independent pathway. *J. Exp. Med.* 187, 237.
- Malvar, T.; Biron, R.W.; Kaback, D.B.; Denis, C.L. (1992). The CCR4 protein from *Saccharomyces cerevisiae* contains a leucine rich repeat region which is required for its control of *ADH2* gene expression. *Genetics*, 132, 951.
- Menezes, J.; Leibold, W.; Klein, G.; Clements, G. (1975). Establishment and characterization of an Epstein-Barr virus (EBV)-negative lymphoblastoid B cell line (BJA-B) from an exceptional, EBV-genome-negative African Burkitt's lymphoma. *Biomedicine* 22, 276.
- Noelle, R.J. (1996). CD40 and its ligand in host defense. *Immunity*, 4, 415-419.
- Puech, A.; Dupressoir, A.; Loireau, M.; Mattei, M.; Heidemann, T. (1997). Characterization of two age-induced intracisternal A-particle-related transcripts in the mouse liver. *J. Biol. Chem.* 272, 5995.

Stamenkovic, I.; Clark, E.A.; Seed, B. (1989). A B-lymphocyte activation molecule related to the nerve growth factor receptor and induced by cytokines in carcinomas. *Embo J.* 8, 1403.

# SEQUENCE LISTING

## (1) GENERAL INFORMATION:

### (i) APPLICANT:

(A) NAME: Vlaams Interuniversitair Instituut voor  
Biotechnologie  
(B) STREET: Rijvisschestraat 120  
(C) CITY: Zwijnaarde  
(E) COUNTRY: Belgium  
(F) POSTAL CODE (ZIP): 9052B  
(G) TELEPHONE: +32 9 2446611  
(H) TELEFAX: +32 9 2446610

(ii) TITLE OF INVENTION: CD40 interacting and TRAF interacting proteins

(iii) NUMBER OF SEQUENCES: 6

### (iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

## (2) INFORMATION FOR SEQ ID NO:1:

### (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1920 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

### (vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo sapiens

### (ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION:20..1108

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GTGCAGAGGC GGCAGGAAGA TGGAGTTGGG GAGTTGCCTG GAGGGCGGGA GGGAGGCGGC	60
GGAGGAAGAG GGCGAGCCTG AGGTGAAAAA GCGGCGACTT CTGTGTGTGG AGTTTGCCTC	120
GGTCGCAAGC TGCGATGCCG CAGTGGCTCA GTGCTTCCTG GCCGAGAACG ACTGGGAGAT	180
GGAAAGGGCT CTGAACTCCT ACTTCGAGCC TCCGGTGGAG GAGAGCGCCT TGGAACGCCG	240

ACCTGAAACC ATCTCTGAGC CCAAGACCTA TGTGACCTA ACCAATGAAG AAACAACCTGA	300
TTCCACCACT TCTAAAATCA GCCCATCTGA AGATACTCAG CAAGAAAATG GCAGCATGTT	360
CTCTCTCATT ACCTGGAATA TTGATGGATT AGATCTAAAC AATCTGTCAG AGAGGGCTCG	420
AGGGGTGTGT TCCTACTTAG CTTTGTACAG CCCAGATGTG ATATTTCTAC AGGAAGTTAT	480
TCCCCCATAT TATAGCTACC TAAAGAAGAG ATCAAGTAAT TATGAGATTA TTACAGGTCA	540
TGAAGAAGGA TATTTTACAG CTATAATGTT GAAGAAATCA AGAGTGAAAT TAAAAAGCCA	600
AGAGATTATT CCTTTTCCAA GTACCAAAAT GATGAGAAAC CTTTTATGTG TGCATGTGAA	660
TGTGTCAGGA AATGAGCTTT GCCTTATGAC ATCCCATTG GAGAGCACCA GAGGGCATGC	720
TGCGGAACGA ATGAATCAGT TAAAAATGGT TTTAAAGAAA ATGCAAGAGG CTCCAGAGTC	780
AGCTACAGTT ATATTTGCAG GAGATACAAA TCTAAGGGAT CGAGAGGTTA CCAGATGTGG	840
TGGTTTACCC AACCAACATTG TGGATGTCTG GGAGTTTTTG GGCAAACCTA AACATTGCCA	900
GTATACATGG GATACACAAA TGAACCTCTAA TCTTGGAATA ACTGCTGCTT GTAAACTTCG	960
TTTTGATCGA ATATTTTTCA GAGCAGCAGC AGAAGAGGGA CACATTATTC CCCGAAGTTT	1020
GGACCTTCTT GGATTAGAAA AACTGGACTG TGGTAGATTT CCTAGTGATC ACTGGGGTCT	1080
TCTGTGCAAC TTAGATATAA TATTGTAAAA TGCTTTTCAA GTGTGGGTTT TGCCCTGATT	1140
GTTGCAAATA CAATTTCCAC CTTCTGGAAA GGTAGGTTTG CTGTGGAGGA AATAATGTAC	1200
TAGATCATTG TCACAGAAAA ACCAACTATG ATTTATGGTT GTGTTTTTCAG AATTCAACAT	1260
TAAAGATTAA TGTTTATTTA AACGAACACA TTCCTGCATT CAGGATGTGA GGCCATTTAA	1320
TAAAAAGGGC ACAAAGCCTG TCAGAGTTTT CAACGGTGCT TACAGCTGCC AGCTGGATTC	1380
CAAACAGGTA CCCCATTGTC TCTGAGCTAA TGTTTATATT TTTCCATTCA GGCACCGAAA	1440
TAGTTAATAT TTAAAATAAG TCTTCAAAAG AAAACATAAG AGATTATTGA GTTCTTGGGA	1500
CTGGATCCTT TATTTCATAA GTTCAGATCA TCTTAAATGA AAATGCCATG ATTATCTGCA	1560
GTTAAGTAGA TGACAGCTAT TCTACATCAG ACTTGATTTT TGTCAGCTAA TTACATAATT	1620
GGTAAGNTAT AATTGAAACC TTATGGCTTA AAATTCCTTA ACTCCTTTTT GATTCATGTT	1680
TGTAGTCATG TTGTCAACAG AGGCAAAGTT AAGCTTGATG ATGGTTAAAA TCGGTTTGAT	1740
AGCACCATGG GACATTTTTT TAACAAAAAT AAATGCATGA AGAGACATAG CCTTTTAGTT	1800
TTGCTAATTG TGAAATGGAA ATGCTTTACA GGAAGTAAAT GCAAATTANT TTTAAGTGTG	1860
CTTTAAAGAA AAATATTTTC CCCACAGGAG AAATTTAAAT AAAGAATTTT ATTTGGTAAA	1920

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 362 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
    (A) ORGANISM: Homo sapiens

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Glu	Leu	Gly	Ser	Cys	Leu	Glu	Gly	Gly	Arg	Glu	Ala	Ala	Glu	Glu	
1				5				10						15		
Glu	Gly	Glu	Pro	Glu	Val	Lys	Lys	Arg	Arg	Leu	Leu	Cys	Val	Glu	Phe	
			20					25					30			
Ala	Ser	Val	Ala	Ser	Cys	Asp	Ala	Ala	Val	Ala	Gln	Cys	Phe	Leu	Ala	
		35				40						45				
Glu	Asn	Asp	Trp	Glu	Met	Glu	Arg	Ala	Leu	Asn	Ser	Tyr	Phe	Glu	Pro	
	50					55					60					
Pro	Val	Glu	Glu	Ser	Ala	Leu	Glu	Arg	Arg	Pro	Glu	Thr	Ile	Ser	Glu	
65				70					75						80	
Pro	Lys	Thr	Tyr	Val	Asp	Leu	Thr	Asn	Glu	Glu	Thr	Thr	Asp	Ser	Thr	
			85					90						95		
Thr	Ser	Lys	Ile	Ser	Pro	Ser	Glu	Asp	Thr	Gln	Gln	Glu	Asn	Gly	Ser	
			100					105					110			
Met	Phe	Ser	Leu	Ile	Thr	Trp	Asn	Ile	Asp	Gly	Leu	Asp	Leu	Asn	Asn	
		115					120					125				
Leu	Ser	Glu	Arg	Ala	Arg	Gly	Val	Cys	Ser	Tyr	Leu	Ala	Leu	Tyr	Ser	
		130				135						140				
Pro	Asp	Val	Ile	Phe	Leu	Gln	Glu	Val	Ile	Pro	Pro	Tyr	Tyr	Ser	Tyr	
145				150						155					160	
Leu	Lys	Lys	Arg	Ser	Ser	Asn	Tyr	Glu	Ile	Ile	Thr	Gly	His	Glu	Glu	
			165					170						175		
Gly	Tyr	Phe	Thr	Ala	Ile	Met	Leu	Lys	Lys	Ser	Arg	Val	Lys	Leu	Lys	
			180					185						190		



Ser	Gln	Glu	Ile	Ile	Pro	Phe	Pro	Ser	Thr	Lys	Met	Met	Arg	Asn	Leu	195	200	205	
Leu	Cys	Val	His	Val	Asn	Val	Ser	Gly	Asn	Glu	Leu	Cys	Leu	Met	Thr	210	215	220	
Ser	His	Leu	Glu	Ser	Thr	Arg	Gly	His	Ala	Ala	Glu	Arg	Met	Asn	Gln	225	230	235	240
Leu	Lys	Met	Val	Leu	Lys	Lys	Met	Gln	Glu	Ala	Pro	Glu	Ser	Ala	Thr	245	250	255	
Val	Ile	Phe	Ala	Gly	Asp	Thr	Asn	Leu	Arg	Asp	Arg	Glu	Val	Thr	Arg	260	265	270	
Cys	Gly	Gly	Leu	Pro	Asn	Asn	Ile	Val	Asp	Val	Trp	Glu	Phe	Leu	Gly	275	280	285	
Lys	Pro	Lys	His	Cys	Gln	Tyr	Thr	Trp	Asp	Thr	Gln	Met	Asn	Ser	Asn	290	295	300	
Leu	Gly	Ile	Thr	Ala	Ala	Cys	Lys	Leu	Arg	Phe	Asp	Arg	Ile	Phe	Phe	305	310	315	320
Arg	Ala	Ala	Ala	Glu	Glu	Gly	His	Ile	Ile	Pro	Arg	Ser	Leu	Asp	Leu	325	330	335	
Leu	Gly	Leu	Glu	Lys	Leu	Asp	Cys	Gly	Arg	Phe	Pro	Ser	Asp	His	Trp	340	345	350	
Gly	Leu	Leu	Cys	Asn	Leu	Asp	Ile	Ile	Leu							355	360		

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1312 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Mus musculus
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION:122..1234
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AGCTATTAAT GATTCTGAATT TATACGACTC ACTATAGGGA ATTTGGCCCT CGAGGCCAAG

60

AATTCGGCAC GAGGGCGGGA AGCAGCGTGA AGAGCGGGTG TTTTGAGGGG ACCCTGCGGC	120
G ATG GCG TCT GGC AGC AGT TCC GAT GCG GCG GAG CCC GCA GGG CCG	166
Met Ala Ser Gly Ser Ser Ser Asp Ala Ala Glu Pro Ala Gly Pro	
1 5 10 15	
GCA GGG CGG GCG GCG TCG GCG CCC GAA GCA GCA CAG GCG GAG GAG GAC	214
Ala Gly Arg Ala Ala Ser Ala Pro Glu Ala Ala Gln Ala Glu Glu Asp	
20 25 30	
CGG GTG AAG AGG CGG CGG CTT CAG TGC CTG GGC TTT GCG TTG GTG GGG	262
Arg Val Lys Arg Arg Arg Leu Gln Cys Leu Gly Phe Ala Leu Val Gly	
35 40 45	
GGA TGC GAC CCC ACG ATG GTC CCC AGC GTC CTG CGG GAG AAC GAC TGG	310
Gly Cys Asp Pro Thr Met Val Pro Ser Val Leu Arg Glu Asn Asp Trp	
50 55 60	
CAG ACG CAG AAA GCC CTG AGC GCC TAC TTC GAG CTG CCA GAG AAC GAC	358
Gln Thr Gln Lys Ala Leu Ser Ala Tyr Phe Glu Leu Pro Glu Asn Asp	
65 70 75	
CAA GGG TGG CCG CGC CAG CCT CCC ACG TCC TTC AAG TCC GAG GCC TAT	406
Gln Gly Trp Pro Arg Gln Pro Pro Thr Ser Phe Lys Ser Glu Ala Tyr	
80 85 90 95	
GTT GAT CTA ACC AAC GAG GAT GCA AAT GAT ACA ACC ATT TTA GAA GCC	454
Val Asp Leu Thr Asn Glu Asp Ala Asn Asp Thr Thr Ile Leu Glu Ala	
100 105 110	
AGT CCA TCT GGA ACT CCT CTA GAA GAT AGC AGC ACT ATT TCT TTC ATT	502
Ser Pro Ser Gly Thr Pro Leu Glu Asp Ser Ser Thr Ile Ser Phe Ile	
115 120 125	
ACC TGG AAT ATT GAT GGA TTA GAT GGA TGC AAT CTG CCC GAG AGG GCT	550
Thr Trp Asn Ile Asp Gly Leu Asp Gly Cys Asn Leu Pro Glu Arg Ala	
130 135 140	
CGA GGG GTG TGT TCC TGC CTA GCT TTG TAT AGT CCA GAT GTG GTA TTT	598
Arg Gly Val Cys Ser Cys Leu Ala Leu Tyr Ser Pro Asp Val Val Phe	
145 150 155	
CTA CAG GAA GTT ATC CCC CCA TAC TGT GCC TAC CTA AAG AAG AGA GCA	646
Leu Gln Glu Val Ile Pro Pro Tyr Cys Ala Tyr Leu Lys Lys Arg Ala	
160 165 170 175	
GCC AGT TAC ACA ATT ATT ACA GGT AAT GAA GAA GGA TAT TTC ACA GCT	694
Ala Ser Tyr Thr Ile Ile Thr Gly Asn Glu Glu Gly Tyr Phe Thr Ala	
180 185 190	
ATA CTA TTG AAG AAA GGA AGA GTG AAA TTT AAA AGT CAG GAG ATT ATT	742
Ile Leu Leu Lys Lys Gly Arg Val Lys Phe Lys Ser Gln Glu Ile Ile	
195 200 205	
CCT TTT CCA AAT ACC AAA ATG ATG AGA AAC CTG CTA TGC GTA AAT GTG	790
Pro Phe Pro Asn Thr Lys Met Met Arg Asn Leu Leu Cys Val Asn Val	
210 215 220	

AGT TTG GGT GGA AAT GAA TTT TGC CTT ATG ACA TCC CAT TTG GAG AGC	838
Ser Leu Gly Gly Asn Glu Phe Cys Leu Met Thr Ser His Leu Glu Ser	
225 230 235	
ACC AGA GAA CAT TCT GCG GAA CGA ATA AGA CAA TTA AAA ACT GTT CTT	886
Thr Arg Glu His Ser Ala Glu Arg Ile Arg Gln Leu Lys Thr Val Leu	
240 245 250 255	
GGA AAA ATG CAA GAG GCT CCA GAT TCA ACC ACG GTT ATA TTT GCA GGA	934
Gly Lys Met Gln Glu Ala Pro Asp Ser Thr Thr Val Ile Phe Ala Gly	
260 265 270	
GAT ACA AAT TTA AGA GAT CAA GAA GTT ATC AAA TGT GGT GGT TTA CCT	982
Asp Thr Asn Leu Arg Asp Gln Glu Val Ile Lys Cys Gly Gly Leu Pro	
275 280 285	
GAC AAC GTT TTT GAT GCC TGG GAA TTT TTA GGC AAA CCT AAA CAT TGC	1030
Asp Asn Val Phe Asp Ala Trp Glu Phe Leu Gly Lys Pro Lys His Cys	
290 295 300	
CAG TAT ACA TGG GAT ACG AAA GCA AAT AAC AAC CTC AGG ATC CCT GCT	1078
Gln Tyr Thr Trp Asp Thr Lys Ala Asn Asn Asn Leu Arg Ile Pro Ala	
305 310 315	
GCT TAT AAG CAT CGT TTT GAT CGA ATA TTT TTC AGA GCA GAA GAG GGG	1126
Ala Tyr Lys His Arg Phe Asp Arg Ile Phe Phe Arg Ala Glu Glu Gly	
320 325 330 335	
CAC CTT ATT CCT CAA AGT TTA GAC CTT GTT GGG TTG GAA AAA CTG GAC	1174
His Leu Ile Pro Gln Ser Leu Asp Leu Val Gly Leu Glu Lys Leu Asp	
340 345 350	
TGT GGT AGA TTT CCG AGT GAT CAC TGG GGG CTC TTG TGC ACC TTG AAT	1222
Cys Gly Arg Phe Pro Ser Asp His Trp Gly Leu Leu Cys Thr Leu Asn	
355 360 365	
GTA GTA TTG TGA AAAGCTTCCC ACTTGCAGCT TTACACGTTT GTTAGCACTA	1274
Val Val Leu *	
370	
GTCTCTGAATT TGTGTAGGTC TCAACCTTTC AGGACATC	1312

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 371 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Ala	Ser	Gly	Ser	Ser	Ser	Asp	Ala	Ala	Glu	Pro	Ala	Gly	Pro	Ala
1				5				10					15		

Gly Arg Ala Ala Ser Ala Pro Glu Ala Ala Gln Ala Glu Glu Asp Arg  
                   20                  25                  30  
 Val Lys Arg Arg Arg Leu Gln Cys Leu Gly Phe Ala Leu Val Gly Gly  
                   35                  40                  45  
 Cys Asp Pro Thr Met Val Pro Ser Val Leu Arg Glu Asn Asp Trp Gln  
           50                  55                  60  
 Thr Gln Lys Ala Leu Ser Ala Tyr Phe Glu Leu Pro Glu Asn Asp Gln  
   65                  70                  75                  80  
 Gly Trp Pro Arg Gln Pro Pro Thr Ser Phe Lys Ser Glu Ala Tyr Val  
                   85                  90                  95  
 Asp Leu Thr Asn Glu Asp Ala Asn Asp Thr Thr Ile Leu Glu Ala Ser  
           100                  105                  110  
 Pro Ser Gly Thr Pro Leu Glu Asp Ser Ser Thr Ile Ser Phe Ile Thr  
           115                  120                  125  
 Trp Asn Ile Asp Gly Leu Asp Gly Cys Asn Leu Pro Glu Arg Ala Arg  
   130                  135                  140  
 Gly Val Cys Ser Cys Leu Ala Leu Tyr Ser Pro Asp Val Val Phe Leu  
  145                  150                  155                  160  
 Gln Glu Val Ile Pro Pro Tyr Cys Ala Tyr Leu Lys Lys Arg Ala Ala  
           165                  170                  175  
 Ser Tyr Thr Ile Ile Thr Gly Asn Glu Glu Gly Tyr Phe Thr Ala Ile  
           180                  185                  190  
 Leu Leu Lys Lys Gly Arg Val Lys Phe Lys Ser Gln Glu Ile Ile Pro  
   195                  200                  205  
 Phe Pro Asn Thr Lys Met Met Arg Asn Leu Leu Cys Val Asn Val Ser  
   210                  215                  220  
 Leu Gly Gly Asn Glu Phe Cys Leu Met Thr Ser His Leu Glu Ser Thr  
  225                  230                  235                  240  
 Arg Glu His Ser Ala Glu Arg Ile Arg Gln Leu Lys Thr Val Leu Gly  
           245                  250                  255  
 Lys Met Gln Glu Ala Pro Asp Ser Thr Thr Val Ile Phe Ala Gly Asp  
           260                  265                  270  
 Thr Asn Leu Arg Asp Gln Glu Val Ile Lys Cys Gly Gly Leu Pro Asp  
   275                  280                  285  
 Asn Val Phe Asp Ala Trp Glu Phe Leu Gly Lys Pro Lys His Cys Gln  
   290                  295                  300  
 Tyr Thr Trp Asp Thr Lys Ala Asn Asn Asn Leu Arg Ile Pro Ala Ala  
  305                  310                  315                  320

Tyr Lys His Arg Phe Asp Arg Ile Phe Phe Arg Ala Glu Glu Gly His  
                     325                    330                    335  
 Leu Ile Pro Gln Ser Leu Asp Leu Val Gly Leu Glu Lys Leu Asp Cys  
                     340                    345                    350  
 Gly Arg Phe Pro Ser Asp His Trp Gly Leu Leu Cys Thr Leu Asn Val  
                     355                    360                    365  
 Val Leu \*  
           370

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:  
     (A) LENGTH: 1536 base pairs  
     (B) TYPE: nucleic acid  
     (C) STRANDEDNESS: double  
     (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:  
     (A) ORGANISM: Homo sapiens

(ix) FEATURE:  
     (A) NAME/KEY: CDS  
     (B) LOCATION: 209..1536

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AGAGAAAGAG GCTCCGGGGA GATAGCGGAC CAGTGAGGGC TGCCCCCTCTT TTGAAGCGGT	60
TTTCGTCTCTT TTCCGCCAGT GGCCTCCCAG CTCACGCAGG GGCGGGTCCC GGTAGCGCGA	120
GGCGGTGCAG GGCGGGAAGG GGAGTGGTGG CGGCTGCGGC AGTAGGGACA GCAGGAGCAG	180
TGGTGCTGTC AGCGCGGCCG TCGGAGACAT GGGAGACCCG GGGTCGGAAA TAATAGAATC	240
TGTCCCTCCA GCTGGCCCTG AGGCATCTGA GTCAACAACG GATGAAAATG AAGACGACAT	300
TCAGTTTGTC AGTGAAGGAC CATCGAGACC TGTTCTTGAA TACATCGATC TGGTCTGTGG	360
TGATGATGAA AACCTAGCG CCTATTATAG TGATATTCTG TTTCTAAAA TGCCAAAACG	420
ACAGGGTGAT TTTTTCATT TTTTAAATAT GAAGAAGGTG AAAACAGACA CAGAAAATAA	480
TGAAGTGAGC AAAAATCACT GCAGATTGTC TAAGGCAAAG GAACCACATT TCGAGTATAT	540
AGAACAACCA ATCATTGAAG AAAAGCCATC ACTTTCATCA AAGAAAGAAA TAGATAATCT	600
TGTGCTTCCA GATTGTTGGA ATGAAAAACA AGCATTTATG TTTACAGAAC AATACAAATG	660

GCTTGAAATA AAAGAAGGTA AATTAGGATG TAAGGATTGT TCAGCAGTTC GGCATTTGGG	720
ATCGAAAGCA GAAAAGCATG TCCATGTGTC CAAGGAATGG ATTGCATATT TAGTAACCCC	780
TAATGGCAGT AATAAACTA CTAGGCAAGC TTCTCTACGA AAAAAAATTA GGAACATGA	840
TGTTTCTAAA GCCCATGGTA AAATTCAGGA TTTGTTAAAG GAATCAACTA ATGATTCAAT	900
TTGTAATTTA GTGCATAAAC AAAATAATAA AAATATTGAT GCTACTGTAA AAGTTTTC	960
TACTGTTTAC AGTTTAGTAA AACATAACAG ACCTTTATCT GATATTGAGG GGGCAAGAGA	1020
ATTACAGGAA AAAAATGGAG AGGTAAATTG TTTAAATACA CGTTACAGTG CAACAAGAAT	1080
AGCAGAACAT ATTGCAAAAAG AAATGAAGAT GAAGATATTT AAGAATATTA TAGAAGAGAA	1140
TGCCAAAATC TGTATCATAA TTGATGAGGC ATCTACAGTT TCAAAGAAAA CCACCCTAGT	1200
GATTTATCTC CAGTGCACAA TTCAGTCAGC TCCTGCACCT GTTATGTTAT TTGTGGCTTT	1260
AAAAGAATTG GTGTCAACTA TAGCAGAGTG TATTGTCAAT ACATTATTGA CTACTTTAAA	1320
TGATTGTGGT TTTACAAATG AATATTTGAA AGCAAATTTA ATTGCATTTT GTTCTGATGG	1380
TGCTAATACA ANCCTGGGAA GAAAGTCTGG AGTAGCTACA AAATTGTTAG AAAATTTTCC	1440
TGAAATCATC ATTTGGAACT GTTTAAATCA TCGATTACAA TTGTCACTTG ATGATTCTAT	1500
ATCCGAAATA AAACAAATTA ATCATTTAAN NTATAA	1536

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 442 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
 

- (A) ORGANISM: homo sapiens

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met	Gly	Asp	Pro	Gly	Ser	Glu	Ile	Ile	Glu	Ser	Val	Pro	Pro	Ala	Gly
1				5					10					15	
Pro	Glu	Ala	Ser	Glu	Ser	Thr	Thr	Asp	Glu	Asn	Glu	Asp	Asp	Ile	Gln
			20					25					30		

Phe Val Ser Glu Gly Pro Ser Arg Pro Val Leu Glu Tyr Ile Asp Leu  
35 40 45  
Val Cys Gly Asp Asp Glu Asn Pro Ser Ala Tyr Tyr Ser Asp Ile Leu  
50 55 60  
Phe Pro Lys Met Pro Lys Arg Gln Gly Asp Phe Leu His Phe Leu Asn  
65 70 75 80  
Met Lys Lys Val Lys Thr Asp Thr Glu Asn Asn Glu Val Ser Lys Asn  
85 90 95  
His Cys Arg Leu Ser Lys Ala Lys Glu Pro His Phe Glu Tyr Ile Glu  
100 105 110  
Gln Pro Ile Ile Glu Glu Lys Pro Ser Leu Ser Ser Lys Lys Glu Ile  
115 120 125  
Asp Asn Leu Val Leu Pro Asp Cys Trp Asn Glu Lys Gln Ala Phe Met  
130 135 140  
Phe Thr Glu Gln Tyr Lys Trp Leu Glu Ile Lys Glu Gly Lys Leu Gly  
145 150 155 160  
Cys Lys Asp Cys Ser Ala Val Arg His Leu Gly Ser Lys Ala Glu Lys  
165 170 175  
His Val His Val Ser Lys Glu Trp Ile Ala Tyr Leu Val Thr Pro Asn  
180 185 190  
Gly Ser Asn Lys Thr Thr Arg Gln Ala Ser Leu Arg Lys Lys Ile Arg  
195 200 205  
Glu His Asp Val Ser Lys Ala His Gly Lys Ile Gln Asp Leu Leu Lys  
210 215 220  
Glu Ser Thr Asn Asp Ser Ile Cys Asn Leu Val His Lys Gln Asn Asn  
225 230 235 240  
Lys Asn Ile Asp Ala Thr Val Lys Val Phe Asn Thr Val Tyr Ser Leu  
245 250 255  
Val Lys His Asn Arg Pro Leu Ser Asp Ile Glu Gly Ala Arg Glu Leu  
260 265 270  
Gln Glu Lys Asn Gly Glu Val Asn Cys Leu Asn Thr Arg Tyr Ser Ala  
275 280 285  
Thr Arg Ile Ala Glu His Ile Ala Lys Glu Met Lys Met Lys Ile Phe  
290 295 300  
Lys Asn Ile Ile Glu Glu Asn Ala Lys Ile Cys Ile Ile Ile Asp Glu  
305 310 315 320  
Ala Ser Thr Val Ser Lys Lys Thr Thr Leu Val Ile Tyr Leu Gln Cys  
325 330 335

Thr	Ile	Gln	Ser	Ala	Pro	Ala	Pro	Val	Met	Leu	Phe	Val	Ala	Leu	Lys	
			340					345					350			
Glu	Leu	Val	Ser	Thr	Ile	Ala	Glu	Cys	Ile	Val	Asn	Thr	Leu	Leu	Thr	
		355					360					365				
Thr	Leu	Asn	Asp	Cys	Gly	Phe	Thr	Asn	Glu	Tyr	Leu	Lys	Ala	Asn	Leu	
	370					375					380					
Ile	Ala	Phe	Cys	Ser	Asp	Gly	Ala	Asn	Thr	Xaa	Leu	Gly	Arg	Lys	Ser	
385					390					395					400	
Gly	Val	Ala	Thr	Lys	Leu	Leu	Glu	Asn	Phe	Pro	Glu	Ile	Ile	Ile	Trp	
				405					410					415		
Asn	Cys	Leu	Asn	His	Arg	Leu	Gln	Leu	Ser	Leu	Asp	Asp	Ser	Ile	Ser	
			420					425					430			
Glu	Ile	Lys	Gln	Ile	Asn	His	Leu	Xaa	Tyr							
		435					440									